# SHAPE MEASURES FOR IMAGE RETRIEVAL

*George Gagaudakis and Paul L. Rosin*

Cardiff University, Department of Computer Science, Newport Road, Cardiff, Wales CF24 3XF, UK
e-mail:{G.Gagaudakis, Paul.Rosin}@cs.cf.ac.uk

**ABSTRACT**

One of the main goals in Content Based Image Retrieval (CBIR) is to incorporate shape into the process in a reliable manner. In order to overcome the difficulties of directly obtaining shape information (in particular avoiding region segmentation) we develop several shape measures that tackle the problem in an indirect manner, requiring only a minimal amount of segmentation. A histogram-based scheme is then used, maintaining low complexity with high efficiency and robustness. The obtained results showed that the synergy of the shape measures worked out providing an improvement over the colour histogram.

## 1. INTRODUCTION

A limitation with the current state of the art Content Based Image Retrieval (CBIR) systems is that they are mainly restricted to the lowest level of query based on primitive image features such as colour and texture whereas ideally they should operate based on the semantic content (e.g. find pictures of steam trains in the country) [3].

The early CBIR systems used techniques such as colour histograms because they were easy to compute, robust, and fairly effective. However, it soon became apparent that it was necessary to incorporate some spatial information into the search. An obvious approach to include shape is to segment the image into regions. It is then straightforward to measure region shape as well as determining spatial interrelationships between regions. The difficulty is that segmentation is inherently such a difficult task that the performance of current algorithms falls far short of being able to provide an adequate input to such schemes [2].

Rather than perform region segmentation some researchers have investigated the use of interest points as a means of localising processing to significant image windows [9]. Still, the difficulty is that corner detection and other interest operators are typically unreliable. This suggests edge detection as a more reliable approach, since it lies somewhere in between regions and points. It does not require a complete partitioning of the image like region segmentation. Only the edges are of interest, and they only cover a fraction (e.g. a tenth) of the image.

Our goal is to incorporate some aspects of shape information into the CBIR process, preferably without explicitly having to extract shapes (i.e. regions) from the image. In the same vein of obtaining some aspect of shape from edges are approaches by Jain and Vailaya [6] and Zhou and Huang [12]. The former histograms the edge orientations. The latter is based on the analogy of filling edge curves with a flow of water. Various features are extracted such as fill time, the number of forks encountered, the number of loops, etc. Our previous work also explored other non-edge based methods of indirect shape measurement.

## 2. METHODS

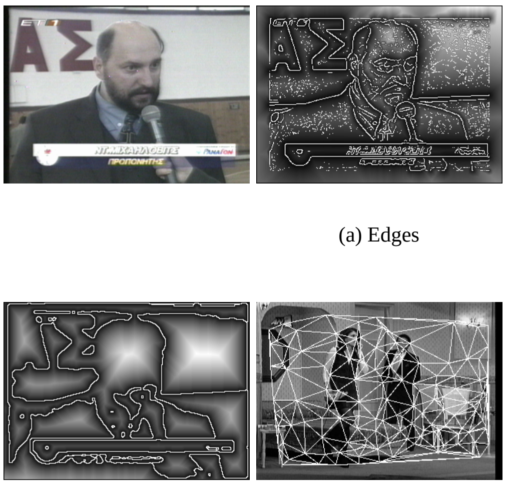
Our objective is to enhance the colour histogram with additional information extracted from images. We exploited our strategy in three steps. We expanded our current methods to take the spatial distribution of colour into account, section 2.1. Then we used statistical methods over local windows to extract aspects of colour and textural characteristics from the image, sections 2.2 and 2.3. Finally we extract different shape aspects in an indirect fashion, section 2.4.

### 2.1. Colour Labels vs. Distance Transforms

In [4] we histogrammed the intensities of the distance map, produced by applying multi-scale distance transform to the detected edges, fig.1. This incorporated spatial information with respect to intensity changes, but ignored colour content. To remedy this we followed a similar trend and applied the multi-scale distance transform to the boundaries of the extracted colour region boundaries, fig.1, creating a 1D histogram of distances to colour edges.

Trying to use colour more actively we expanded the process by combining the colour and shape information, forming a two dimensional histogram. Each bin in that histogram represents the frequency of occurrence of a colour at some distance from a feature. We applied this approach both on detected edges and colour regions boundaries to obtain two histogram features.

A straightforward approach was taken to the colour region segmentation. The objective was not to provide semantically meaningful regions, but rather to provide a stable partitioning. Assuming no severe illumination change, we achieved that by a transformation of the image from the RGB space to a set of perceptual colour labels [1]. To reduce noise we first apply Gaussian smoothing ($\sigma = 2.8$) to the RGB image and after labeling run majority voting filtering using a $5 \times 5$ mask.



(a) Edges

(b) Colour Regions (c) Triangulation

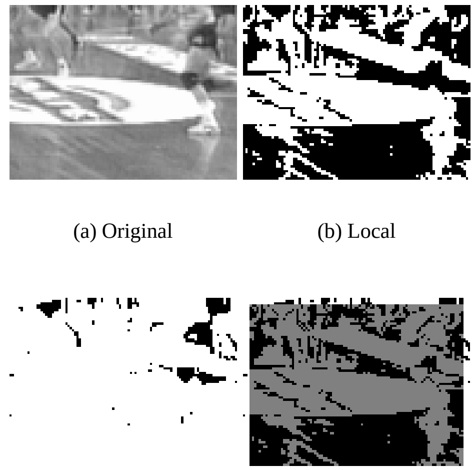
**Fig. 1**. Example of image and the resulting distance transforms with the relevant overlayed colour region boundaries / edges and Triangulation of sub-sampled edges.

### 2.2. Local Colour Entropy

In an attempt to involve more colour in the feature histograms we calculated entropy in local windows over the hue channel of the image. We define a set of $N \times N$ overlapped windows. For each window, centered at $(x, y)$, we generate the histogram of the hues and calculate the entropy ($E_{H_{xy}} = -\sum p_i \log_2 p_i$). The image signature is obtained by histogramming the entropy values over all the windows.

### 2.3. Local vs. Global Statistics

Next we experimented with histogramming the relation between local statistical image information and the corresponding global image information. In particular we wish to capture the similarity and difference of the local statistics of small windows against the corresponding statistics of the whole image. This way we managed to capture the homogeneity of the image as well as the variation of the statistics across the image. This can be considered as capturing some sort of global texture of the image. Thresholding [11] is applied both globally to the image and locally to the individual windows. Then the percentage *difference* between the window and the image content, at the relevant position, is histogrammed. Additionally, we histogrammed the amount of *blackness* fig.2 found in the window content after thresholding. For our experiments we used an overlap of 25% of the window size in each direction.



(a) Original (b) Local

(c) Global (d) Difference

**Fig. 2**. Thresholding of a window(a) and the difference(d) from the patch(c) of the thresholded image..

### 2.4. Delaunay triangulation

In this section we consider another approach to indirectly measuring shape. Its basis is the generation of the Delaunay triangulation [8] of a set of edges. Edge detection and linking is first carried out, to eliminate spurious short edge lists. The remainder is further sub-sampled which both speeds up triangulation and effectively performs some noise suppression. The strength of this approach is that connectivity is used to help filter out noise but nevertheless the triangulation is not dependent on connectivity and therefore can cope with edge linking errors. Unlike Tao and Grosky [10] our scheme does not require isolated objects.

We calculate different aspects of the individual triangles, *Area*, *Aspect ratio* and *Length* and histogram their properties over the complete triangulation. In fig.1, an example of an image and the obtained triangulation is illustrated.

## 3. TESTING PROTOCOL

Our objective is to improve the effectiveness of the simple colour histogram by incorporating shape. At this point we examine the performance and potential of these shape methods when combined with other features. Two main aspects are of interest in this context, namely, *effectiveness* and *efficiency* [4].

To avoid the difficulty (and impracticality) of manual groundtruthing we follow a method similar to Milanese *et al*. [7] where still images extracted from video clips are used. A broadcast TV signal from a local Greek station (CRETA Channel) was captured at a resolution of two frames-per-second. This was then re-sampled to obtain 9 still images representing each clip making a total of about 400 images. As in [7] we assumed that (a) the continuity of the visual content of a clip is implied by the uninterrupted recording of a video camera, (b) there is gradual change of content, from frame to frame, due to camera operations and subject motion, object appearance and disappearance.

## 4. RESULTS

In this paper we investigate the potential of shape measures in image retrieval, when used individually or combined with other measures. Fourteen methods were considered, including those described in this paper as well as in [4], table1 shows the individual methods and ranking. We developed a system to test all the possible combinations of those 14 methods, giving a total of over 16000 combinations.

| Method | Description | Recall | Precision |
|---|---|---|---|
| TXTR | Circular Co-occurrence Matrix | 83.49 | 8.58 |
| TARE | Triangle Areas | 42.63 | 28.22 |
| CL01 | Spatialised Colour Labels | 82.77 | 9.07 |
| LBLK | Local Blackness | 38.73 | 39.54 |
| CL | Berlin & Kay Labels | 82.51 | 6.95 |
| DIFF | Local Binary Difference | 32.27 | 44.24 |
| BD2D | Colour Labels vs. MSDT of Colour Region Boundaries | 81.96 | 10.40 |
| TLEN | Triangulation Edge Length | 32.21 | 40.48 |
| ORNT | Edge Orientation | 76.71 | 16.10 |
| BKDT | MSDT of Colour Region Boundaries | 30.86 | 58.46 |
| CLSH | Colour Labels vs. MSDT of extracted Edges | 73.87 | 13.26 |
| DST | MSDT of extracted Edges | 22.83 | 61.96 |
| ENTR | Local Hue Entropy | 72.75 | 11.16 |
| TRAT | Ratio of Triangles | 19.93 | 61.72 |

**Table 1**. Methods, Description and individual performance.

Brief investigation showed that some methods seem to be appearing in most (if not all) the combinations. Particularly, the *hue entropy* and the *texture* methods are on the top of the list. Followed by the *edge orientation* and the *colour labels vs. MSDT of colour region boundaries*. The methods using the triangulationbehavior, although not on the top of the list, are involved in the top combinations. That shows we cannot really predict the behavior of a combination based only on individual histogram performance.

The nature of the image source suggested that noise levels depend on signal strength and external interference which vary in time. Our target was to investigate any potential correlation between noise levels and feature performance. We estimated the noise level of the images by measuring the standard deviation of their grayscale map using a fast noise variance estimation [5] method. In Fig.3a, the noise level (solid line) across the dataset is illustrated overlaid by the performance of a feature (dotted line). These graphs showed that there is no visible dependency of the feature performance to noise.

The results suggested that the clusters may not disjoint as we were expecting. This warned us to take a look at the database from two different aspects. At a individual image level, inspecting the histograms of the images and at a cluster level, inspecting the distances between images.

At an individual image level we created images, that resemble a form of spectrogram, of the image histograms, Fig.3b. Each pixel column on these images is a histogram of an image. The white lines are used to separate the image clusters. The images are post-processed for viewing purposes. This representation enabled us to have a visual feel of the similarity or dissimilarity of the image features, individually. Additionally we derived conclusions on the applicability and discrimination potential of features:

*Global Level Uniformity.* In some cases the features showed some uniformity limiting the range of the histograms.

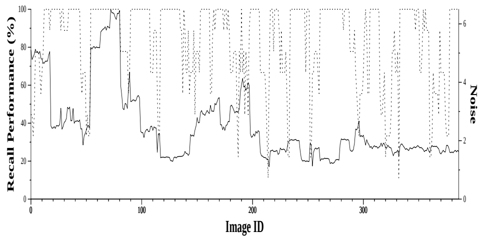
*Noise*. A “noisy” nature of the histogram caused problems, limiting the discrimination power even more.

*Dithering Effect.* Some histograms showed a high degree of uniformity in cluster level and variation in global level, which would be the ideal behavior of a feature. Although, performance was not as good as expected due to small variations of the histograms, it looks like the histograms are shifted slightly, which is not taken into account in the distance measure calculation.

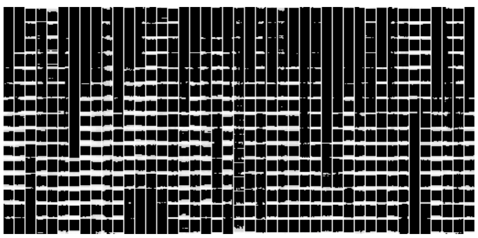
*Narrow Band Bin-Population.* In many cases a narrow band of the histograms is populated.

*Global Level Variation.* This is the kind of behavior of the top performing features. Narrow banding appears in some cases without considerably affecting the discrimination power of the histogram. Such a phenomenon would suggest that compression could be easily achieved by carefully tuning the size of the histogram bins.

At a cluster level, we extracted the distances of all the images to all the images of the database. Since the number of possible combinations is large, we selectively used some features, according to their performance. We used the extracted distances to produce images, Fig.3(c-f) where the position of each pixel represents two image indices and the intensity their distance, the higher the intensity the closer the distance. The matrix is arranged in such a way that inter-cluster distances are represented by $9 \times 9$ pixel blocks on




(a) Texture



(b) TXTR

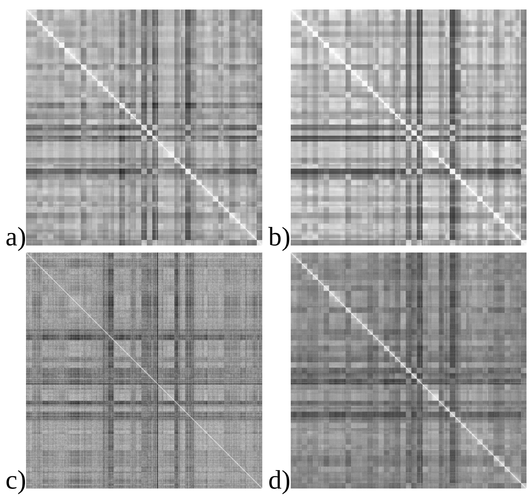



**Fig. 3**. Noise, Histogram Spectrums and Distance matrices.

the diagonal. We used the distance matrices as an indicator for (a)*precision*, nice square blocks on the diagonal and rest of image (b)*recall*, the diagonal is highlighted and (c)*range* of method, high overall contrast. Low performing methods showed negative behavior on all these aspects.

## 5. CONCLUSIONS

This paper describes a number of new shape measures for use in CBIR. The general strategy was to avoid performing region segmentation as this was considered too unreliable. Instead a variety of schemes based on edges were developed. We ran the methods on our existing system and tested all possible combinations of our new and old measures. In many cases by incorporating shape the performance was improved over the plain colour labels histogram. Initial experiment showed that the edge curvature and the colour region shape measure did not work well, and so they were not investigated further. Focusing on the shape aspect, we identified the potential of measuring indirect shape using the Delaunay triangulation.

## 6. REFERENCES


[1] B. Berlin and P. Kay. *Basic Color Terms: their universality and evolution*. University of California Press, 1969.

[2] M.C. Cooper. The tractability of segmentation and scene analysis. *IJCV*, 30(1):27–42, 1998.

[3] J.P. Eakins. Automatic image content retrieval – are we getting anywhere? In *Proc. of Third International Conference on Electronic Library and Visual Information (ELVIRA3)*, pages 123–135, 1996.

[4] G. Gagaudakis and P. L. Rosin. Incorporating shape into histograms for CBIR. *Pattern Recognition*, Forthcoming.

[5] J. Immerkaer. Fast noise variance-estimation. *CVIU*, 64(2):300–302, September 1996.

[6] A.K. Jain and A. Vailaya. Image retrieval using color and shape. *Pattern Recognition*, 29(8):1233–1244, 1996.

[7] R. Milanese and M. Cherbuliez. A rotation, translation, and scale-invariant approach to content-based image retrieval. *Journal of Visual Communication and Image Representation*, 10:186–196, 1999.

[8] F.P. Preparata and M.I. Shamos. *Computational Geometry*. Springer-Verlag, 1985.

[9] C. Schmid and R. Mohr. Local grayvalue invariants for image retrieval. *PAMI*, 19(5):530–535, 1997.

[10] Y. Tao and W. Grosky. Delaunay triangulation for image object indexing: A novel method for shape representation. In *IST SPIE Symposium on Storage and Retrieval for Image and Video Databases VII*, 1999.

[11] W.H. Tsai. Moment-preserving thresholding. *Computer Vision, Graphics and Image Processing*, 29:377–393, 1985.

[12] X.S. Zhou and T.S. Huang. Image representation and retrieval using structural features. In *ICPR00*, pages Vol I: 1039–1042, 2000.